

# H.264 동영상 표준 부호화 방식을 위한 순차적 움직임 벡터 오류 은닉 기법

정종우, 김재훈, 홍민철  
송실대학교 정보통신전자 공학부  
nobreaky@vipl.ssu.ac.kr

## Sequential Motion Vector Error Concealment for H.264 Video Coding

Jong-Woo Jung, Jae-Hoon Kim, Min-Cheol Hong  
Dept. of Electronic Engineering, Soongsil University

### 요 약

본 논문에서는 H.264 표준 동영상 부호화 방식을 위한 순차적 움직임 벡터 오류 은닉 기법을 제안한다. H.264 표준 동영상 부호화 방식에서의 움직임 예측과정이 다양한 크기의 서브 매크로 블록 모드에 따라 각기 다른 움직임 벡터 개수를 갖게 되므로 움직임 벡터는 기존의 표준 부호화 방식에 비해 상대적으로 적은 영역을 대표하게 된다. 그러므로 이웃한 블록의 움직임 벡터간의 상관관계는 서브 매크로 블록의 크기가 작을수록 더 커지게 된다. 변화된 국부 통계 특성에 대한 적응도는  $\alpha$ -trimmed mean 필터를 이용한 부호기의 부호화 순서를 따르는 순차적 움직임 벡터 오류 은닉기법의 성능을 좌우하는 가장 중요한 부분이다. 실험 결과를 통해 제안한 방식이 실시간 동영상 전송에 적합하며 기존 방식과 유사한 성능을 보임을 확인할 수 있었다.

### I. 서론

동영상 전송에 있어 주된 문제점은 압축된 동영상의 비트스트림이 때때로 채널 오류에 의해 손상된다는 것이다. 압축된 동영상 비트스트림의 손상은 인트라 프레임의 손상과 인터프레임의 손상, 또한 압축 정보의 손상으로 구분할 수 있다. 상기 문제점을 해결하기 위해 동영상 압축 방식의 발전에 따라 각 동영상 표준 압축방식에 따른 매우 많은 효과적인 오류 은닉기법이 제안되어 왔다. 그러나 새로운 동영상 표준 부호화 방식인 H.264는 기존의 H.263이나 MPEG4 part 2와 같은 동영상 압축 방식 대비 여러 새로운 압축 기법을 채택하고 있다 [1]. 그러므로 H.264 동영상 표준 압축방식에 적합한 오류 은닉 기법이 요구되고 있으며 본 논문에서는 H.264를 위한 효과적인 움직임 벡터 복원 방식에 대해 기술한다.

H.264 동영상 표준 압축방식의 가장 주된 변화는 움직임 예측 방식의 변화이며, H.264에서는  $16 \times 16$  매크로블록 단위의 움직임 예측이 아닌 하나의 매크로 블록을 여러 개의 다양한 크기의 서브 매크로블록으로 나누어 각각의 움직임 예측을 수행함으로써 압축효율을 높인다. 그러므로

각각의 서브 매크로블록은 고유의 움직임 벡터를 갖게 되어 하나의 매크로 블록의 움직임 벡터의 수는 움직임 벡터 예측을 위한 서브 매크로블록 모드에 의해 결정된다. 때문에 기존의 블록 경계 매칭 기법이나 중첩 움직임 보상 방식들은 H.264에서는 서브 블록의 움직임 예측 특성에 적합하지 못하다 [2][3]. 또한 기존의 H.264를 위한 움직임 벡터의 오류 은닉 기법인 수정된 블록 경계 매칭 방식은 복호기에서 움직임 예측을 수행해야 하는 계산량과 복잡도 문제가 있으며, 라그랑지 보간법을 이용한 방식과 함께 동영상 압축방식의 부호화 순서가 아닌 블록 인터리빙을 가정한다 [4],[5]. 블록 인터리빙을 적용하는 경우는 부호화 순서의 변화로 인한 부호화 버퍼링 지연과 복호화 지연 및 복호기에서의 인터리빙 지원 여부등의 여러 가지 문제점을 갖고 있기 때문에 H.264를 위한 부호기의 부호화 순서에 부합하는 움직임 벡터의 오류 은닉기법이 요구된다.

본 논문에서는 H.264 동영상 표준 부호화 방식의 순차적인 부호화 순서를 따르며 비교적 적은 계산량을 가지는 ATM(alpha-trimmed mean) 필터를 이용한 움직임 벡터 복원 알고리즘에 대해 제안한다.

## II. ATM(Alpha trimmed mean) 필터

입력 신호  $\{x(n)\}$ 에 대한 ATM 필터의 출력  $\{y(n)\}$ 은 아래 식(1)과 같다 [6].

$$y(n) = \left\{ \sum_{j=\hat{\alpha}N+1}^{N-\hat{\alpha}N} x_{(j/N)}(n) \right\} / \{N(1-2\hat{\alpha})\} \quad (1)$$

여기서  $x_{(j/N)}(n)$ 은  $N=2M+1$  개의 입력 신호를 크기순으로 재 정렬한 것이다. 즉,  $x_{(1/M)}(n) \leq \dots \leq x_{(N/M)}(n)$ 의 방식으로 재 정렬한 것이다. 여기서  $\alpha$ 를  $\hat{\alpha} = \{(N-1)/2M\} - \alpha$ 라고 정의할 수 있다면  $y(n)$ 은 아래 식(2)와 같다.

$$y(n) = \left\{ \sum_{j=N_M-\alpha N}^{N_M+\alpha N} x_{(j/N)}(n) \right\} / (2\alpha N + 1) \quad (2)$$

where

$$N_M = (N+1)/2$$

여기서 파라미터  $\alpha$ 는 연산되는 입력 신호의 개수를 결정하는 변수로 ATM(alpha trimmed mean) 필터는 데이터 특성에 따르는 사전정보를 이용한 고정된  $\alpha$ 값에 따라, 또는 국부적인 데이터의 통계적인 특성에 따르는 가변의  $\alpha$ 값에 따라 조절되는 미디언 데이터를 원점으로 하는 필터 탭 내 데이터들의 평균을 취하는 필터이다. 본 논문에서 제안 하는 움직임 벡터의 오류 은닉 기법에서 사용하는 ATM 필터는 필터 탭의 크기를 조절하는  $\alpha$ 를 이용하는 대신에 마스크 내의 데이터 분포에 따른 분산을 이용하여 윈도우의 활동성을 이용하여 필터 탭을 조절 하는 수정된 ATM 필터를 사용한다.

## III. ATM(alpha trimmed mean) 필터를 위한 필터 윈도우의 구성

H.264 에서는 아래 그림1 에서와 같이 움직임 예측 과정에서 하나의 매크로 블록이 여러 개의 다른 크기의 서브매크로 블록으로 나뉘지게 되고 따라서 움직임 벡터의 개수는 1에서 16의 범위를 갖게 된다. 또한 실제 JM(joint model)에서 각각의 서브매크로 블록의 움직임 벡터는 4x4 블록 단위로 모두 16개를 갖게 된다. 예를 들어 8x4 블록의 움직임 벡터는 2개가 존재하게 된다. 따라서 하나의 매크로 블록에 대한 움직임 벡터의 배열 개수는 x, y 방향 모두 16개가 된다. 따라서 본 논문에서 움직임 벡터의 오류 은닉에 대한 기본 단위는 4x4 블록이 된다.

순차적 복원의 ATM 필터를 위한 이웃한 4x4 서브 매크로 블록의 움직임 벡터들로 구성된 필터의 윈도우는 마코프(Markov) 순서를 다르며 아래 그림2 와 같이 나타내

고 X, Y 윈도우는 아래 식(3)과 같이 표현된다.

$$X(i) = MV_x(n-i), \quad Y(i) = MV_y(n-i) \quad (3)$$

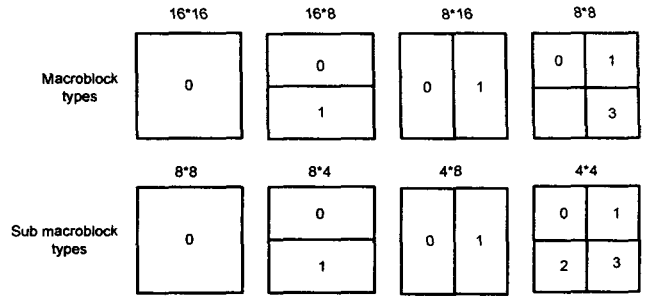


그림 1. ME를 위한 VBS(variable block size) 분할 모드

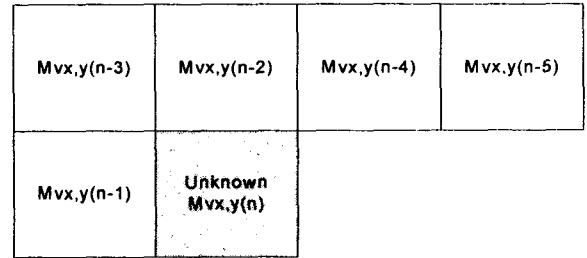


그림 2. ATM 필터의 데이터 윈도우

여기서  $j$ 는 윈도우의 인덱스를 나타낸다. 손상된 하나의 매크로 블록에서 복원될 16개의 움직임 벡터를 위한 4x4 단위의 서브매크로 블록을 그림3 에서와 같이 3개의 영역으로 나누어 필터에 필요한 마스크를 정의 한다. 영역1의 경우는 순차적 복원에 따르는 필터 윈도우 내의 이웃한 움직임 벡터들이 모두 유효한 영역이며 영역2와 영역3은 각각 1개와 2개의 이웃한 움직임 벡터가 유효하지 않은 영역이다. 따라서 그림4 에서와 같이 영역에 따른 필터 윈도우를 정의할 수 있다.

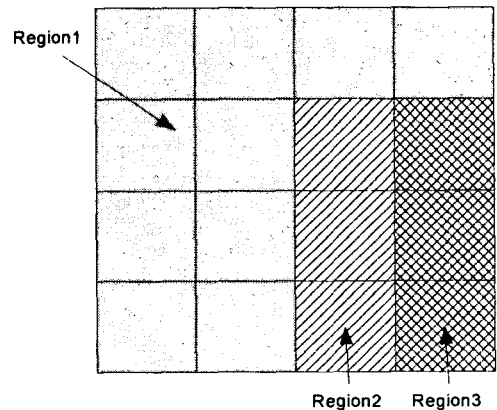


그림 3. 손상된 매크로 블록의 영역 분류

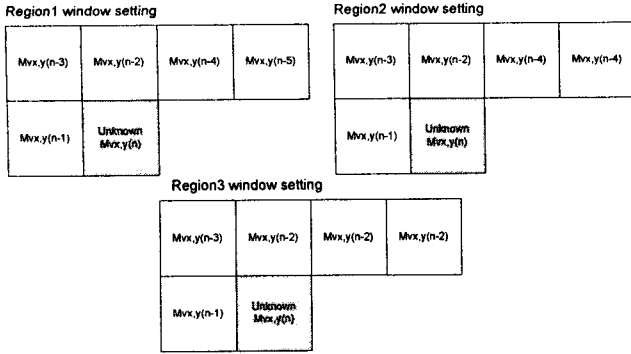


그림 4. 영역별 윈도우 구성

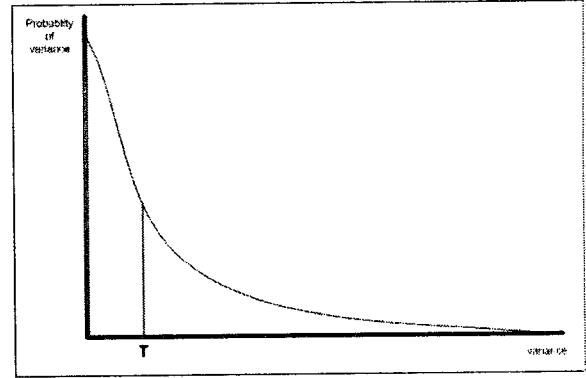


그림 5. 데이터 윈도우 분산 값의 확률 분포

#### IV. 국부통계특성 기반의 ATM(alpha trimmed mean) 필터

움직임 벡터의 오류 은닉 기법에 있어 이웃한 움직임 벡터의 국부 통계 특성에 대한 적응성은 성능을 좌우하게 된다. 3장에서 구성한 x, y 방향의 데이터 윈도우를 이용하여 각 방향에 대한 국부 통계 특성인 평균과 분산을 다음 식(4)와 같이 정의 할 수 있다.

$$Mean(n) = \left\{ \sum_{i=1}^N x_{(i/M)}(n) \right\} / N \quad (4)$$

$$Var(n) = \left\{ \sum_{i=1}^N |x_{(i/M)}(n) - Mean(n)| \right\} / N$$

여기서 국부통계특성인 분산은 계산량의 절감을 위해 절대값으로 근사화 하였다.

식(4)에서 정의한 국부통계특성을 이용하여 데이터 윈도우의 평탄함의 기준이 되는 윈도우의 활동성 ACT의 수식은 다음 식(5)와 같다.

$$Act_X(n) = \sigma_X^2(n) / \{ \sigma_X^2(n) + T_X \} \quad (5)$$

$$Act_Y(n) = \sigma_Y^2(n) / \{ \sigma_Y^2(n) + T_Y \}$$

where

여기서 파라미터  $T$ 는 데이터 윈도우의 들연한 변화의 기준이 되는 분산 값이다.

파라미터  $T$ 의 결정을 위해 다양한 동영상 시퀀스에 대해 x, y 방향의 데이터 윈도우의 120만개의 분산값을 이용하여 아래 그림5와 같이 데이터 윈도우 국부통계특성의 확률 분포를 획득하였다. 파라미터  $T$ 는 누적 확률 값이 1/3 이 되는 지점으로 설정하였다.

데이터 윈도우의 변화 정도를 나타내는 ACT를 이용하여 최종 파라미터  $\alpha$ 는 다음 식(6)과 같이 결정할 수 있다. 식(6)에서와 같이 윈도우의 활동성에 따라  $\alpha$ 는 0에서 M까지의 변화 정도를 가지며 그에 따라 필터는 Mean에서 Median의 형태로 적용하게 된다.

$$\alpha_X(n) = \left[ (1 - Act_X(n))M \right] \quad (6)$$

$$\alpha_Y(n) = \left[ (1 - Act_Y(n))M \right]$$

where  
 $[\cdot]$  is rounding function  
 $N = 2M + 1$

마지막으로 최종 복원되는 움직임 벡터는 다음 식(7)과 같다.

$$\widehat{MV}_X(n) = \sum_{i=N_M - \alpha_X(n)}^{N_M + \alpha_X(n)} X(i) \quad (7)$$

$$\widehat{MV}_Y(n) = \sum_{i=N_M - \alpha_Y(n)}^{N_M + \alpha_Y(n)} Y(i)$$

#### V. 실험 결과

본 논문에서 제안한 순차적 움직임 벡터 오류 은닉 방식은 시간적 대체 방식, Mean 필터, Median 필터 그리고 블록 경계 영역 정합 방식과의 객관적 주관적 성능을 비교하기 위해 QCIF foreman, mobile, news 동영상을 다양한 QP에 대해 실험 하였다.

실험에 사용한 영상은 JVT JM 2.1을 이용하여 부호화, 복호화 하였으며 10frame/s로 부호화 하였다. 제안된 방식의 성능을 측정하기 위해, PSNR이 사용되었으며, 이는 화소 당 8비트로 구성된 M×N 영상 크기에 대해 다음과 같이 정의된다.

$$PSNR = 10 \log \frac{MN \times 255^2}{\|f - \hat{f}\|} \quad (8)$$

식(8)의  $\| \cdot \|$ 은 Euclidean norm 을 의미하며,  $f$  및  $\hat{f}$  은 각각 원 영상과 복원 영상을 의미한다.

블록 경계 영역 정합 방식은 순차적 복원 방식에 따라 실험 하였다. 또한 실험을 위해 하나의 프레임 내의 33%의 매크로 블록이 손실된 경우의 에러 환경을 고려하였다. 실험 결과는 표1과 같다.

video sequence	QP	original PSNR(dB)	simulation result	PSNR of 33% Macroblock lost rate				
				TR	BBM	A-Trimmed	Mean	Median
foreman	16	36.6345	PSNR	30.76	31.27	32.57	32.14	32.34
			MV Error	319382	3028125	2229375	214875	262625
	20	34.1459	PSNR	29.62	29.85	31.14	31.16	31
			MV Error	4669753	291875	2775281	3136938	2680618
	24	31.5203	PSNR	28.51	29.01	29.88	29.62	29.57
			MV Error	4206333	304791675	1872466	1668243	1949324
28	28.7811	PSNR	26.59	27.24	27.99	27.21	27.25	
		MV Error	4930768	354375	2974026	338961	2965097	
mobile	16	33.77	PSNR	27.32	29.5	30.46	29.53	29.68
			MV Error	1226042	113229175	0757324	0911621	0756348
	20	30.6	PSNR	25.58	27.54	29.62	27	27.59
			MV Error	0715104	027395825	0216328	0358398	0227539
	24	27.56	PSNR	23.8	25.7	27.01	25.63	25.92
			MV Error	0770474	043854175	0324597	0378024	0300403
28	24.69	PSNR	22.24	23.47	24.39	23.72	23.81	
		MV Error	0833333	04	0407246	0443359	040332	
news	16	37.1915	PSNR	32.11	32.55	32.84	32.85	32.85
			MV Error	1409226	1395	1244565	1247283	1293478
	20	34.398	PSNR	30.23	30.32	30.64	30.36	30.2
			MV Error	2087708	3133333	2082589	2048875	1868304
	24	31.4458	PSNR	28.35	28.39	29.1	28.41	28.29
			MV Error	18575	2085417	19575	1.97	1.965
28	28.6003	PSNR	27.13	26.9	27.23	27.2	27.13	
		MV Error	2855769	2983333	2889423	2788462	2975962	

표1. 실험 결과

표1에서 알 수 있듯이 제안한 방식은 Mean과 Median보다 더 좋은 성능을 보이거나 둘 중에 더 좋은 쪽의 성능에 가까운 성능을 보인다. 또한 움직임 벡터 에러는 손상된 블록 당 복원한 움직임 벡터와 원 움직임 벡터의 차이 값으로 PSNR과 거의 동일한 성능을 보인다.

주관적인 화질 비교는 그림6 에서 foreman QCIF 영상에 대한 실험 결과를 통해 확인할 수 있다.

## VI. 결론

본 논문에서는 이웃한 움직임 벡터의 국부 통계 특성에 따라 마스크의 크기가 결정되는 ATM 필터를 이용하는 순차적 움직임 벡터 오류 은닉 기법을 제안 하였다. 실험 결과에서 볼 수 있듯이 제안된 방식은 H.264 표준 동영상 부호화 방식의 움직임 예측을 위한 다양한 크기의 서브 매크로블록에 적합하며 표준 부호화 방식의 부호화 및 복호화 순서를 따르는 이유로 인터리빙을 사용하는 기존 방식들의 버퍼링 지연에 대한 문제점을 극복하면서 일정한 성능을 유지함을 확인할 수 있었다.

현재는 보다 향상된 성능을 위해 ATM 필터의 국부 통계 특성 적용성을 결정하는 최적의 파라미터  $\alpha$ 의 결정을 위해 보다 데이터 윈도우 분산 값의 보다 정확한 확률 분포 획득과 확률 모델에 따른 분석의 연구가 진행 중에 있다.

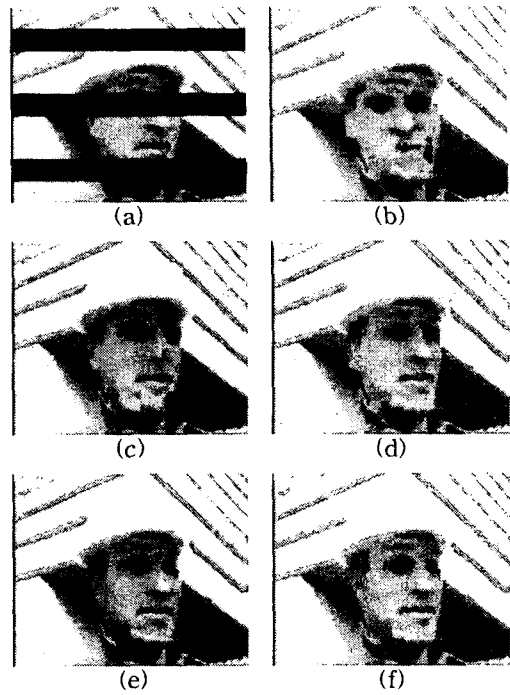


그림 6. Foreman QCIF 영상의 주관적 화질 비교:  
(a)Non-Error concealment, (b)Temporal Replacement  
(c)Block Boundary Matching (d)Mean (e)Median  
(f)Alpha Trimmed Mean

## 참고문헌

- [1] "Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification" Geneva, May 2003.
- [2] M. J. Chen, L. G. Chen, and R. M. Wang, "Error concealment of lost motion vectors with overlapped motion compensation" IEEE Trans. Circuits and Systems for Video Technology, vol. 7, pp. 560-563, June 1997.
- [3] M.E. Al-Mualla, N. Canagarajah, D.R. Bull, "Error Concealment Using Motion Field Interpolation," ICIP '98, vol. 3, Oct 1998.
- [4] Y. Wang, M.M. Hannuksela, V. Varsa, A. Hourunranta, and M. Gabbouj, "The error concealment feature in the H.26L test model," in Proc. Int. Conf. Image Processing 2002, vol. w, Sept. 2002.
- [5] J. H. Zheng, L. P. Chau "a motion vector recovery algorithm for digital video using lagrange interpolation", IEEE Trans. Broadcasting vol. 49, no. 4, Dec. 2003.
- [6] A. Taguchi, "data-dependent a-trimmed mean filters for image restoration" Sep. 1994.